

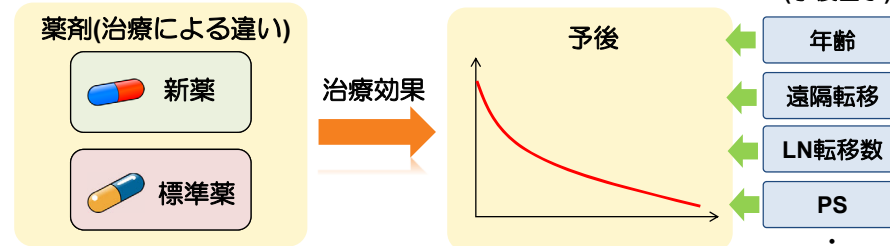
# 医学統計セミナー

## - 生存時間解析(2) - Cox比例ハザードモデル

下川敏雄

和歌山県立医科大学附属病院 臨床研究センター

### 比例ハザードモデルの動機(1)

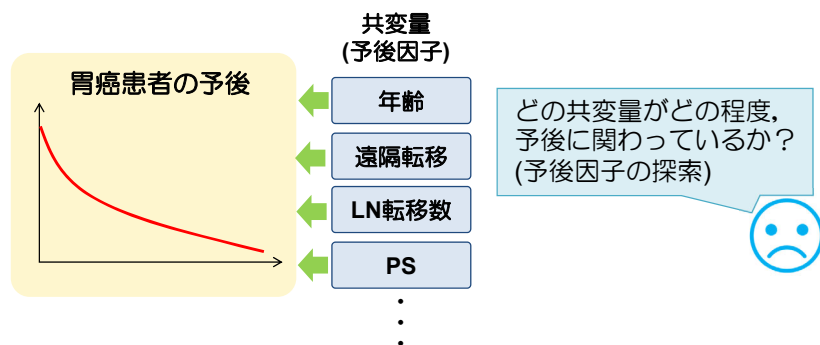


臨床研究では、予後(アウトカム)に影響を及ぼすのが治療(ここでは薬剤)だけではない(治療前の患者の状態(患者背景)がアウトカムに影響を及ぼす)。

例えば、新薬を服薬した群の80%がPS0の患者であり、標準薬を服薬した群の80%がPS2の患者だった場合、たとえ、新薬が標準薬よりも有効性が認められても、

「患者の全身状態が影響したのか、薬が影響したのかを評価できない」  
これら共変量の影響を調整したうえで治療効果を評価する場合に、多変量解析が用いられる。このとき、生存時間データを解析する場合の多変量解析手法が、Cox比例ハザードモデルである。

### 比例ハザードモデルの動機(2)



観察研究では、患者の予後に影響を及ぼす要因を探索する研究が行われる。そこでは、それぞれの共変量が予後に対してどの程度影響しているのかをオッズ比で評価することが多い。

このときに、比例ハザードモデルを用いることで、他の共変量の影響を調整しながら、それぞれの共変量の影響を評価することができる。

### Cox比例ハザードモデル

- モデルの記述： $h(t, \mathbf{x}) = h_0(t) \exp\{\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p\} = h_0(t) \exp\{\boldsymbol{\beta}^T \mathbf{x}\}$
- 記法
- $h_0(t)$ : 基線ハザード(Baseline hazard)
  - 共変量に依存しない潜在的なハザード
  - $\mathbf{x}$ :  $p \times 1$  共変量ベクトル
  - $\boldsymbol{\beta}$ :  $p \times 1$  パラメータ・ベクトル

モデルを書き直せば、

$$\log\{h(t, \mathbf{x})\} = h_0(t) \log\left[\exp\{\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p\}\right] = \log\{h_0(t)\} + \boldsymbol{\beta}^T \mathbf{x}$$

となり、一種の回帰モデルであることがわかる。さらに

$$\log\left\{\frac{h(t, \mathbf{x})}{h_0(t)}\right\} = \boldsymbol{\beta}^T \mathbf{x}$$

と表すこともできる。すなわち、重回帰モデルは、共変量の有無によるハザードの比(ハザード比)の対数値を表している。

- 比例ハザードモデルによる生存率の記述：

$$S(t, \mathbf{x}) = S_0(t) \exp\{-\beta_1 x_1 - \beta_2 x_2 - \dots - \beta_p x_p\} = S_0(t) \exp\{-\boldsymbol{\beta}^T \mathbf{x}\}$$

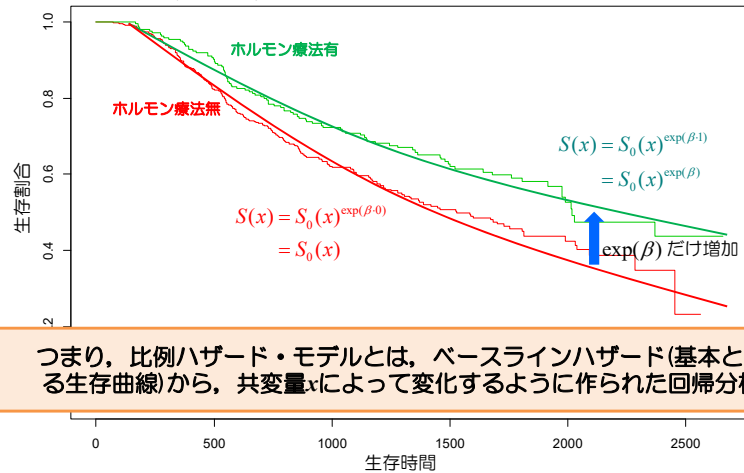
生存時間関数に基づいて比例ハザードモデルを眺める(次ページ)

## Cox比例ハザードモデルの意味

$$\text{モデル: } h(x) = h_0(x) \exp(\beta x)$$

ドイツで実施された、ホルモン療法の有無による有効性・安全性を評価するための無作為化比較第III相試験の例

ホルモン療法有りを1, 無しを0とする



つまり、比例ハザード・モデルとは、ベースラインハザード(基本となる生存曲線)から、共変量xによって変化するように作られた回帰分析

## 比例ハザードモデルにおける「比例ハザード性」

比例ハザードモデル

$$h(t; x) = h_0(t) \exp(\beta_1 x_1 + \beta^T x)$$

$x_1$ が $z_1, z_2$ をとるときのモデル

$$h(t; z_1) = h_0(t) \exp(\beta_1 z_1 + \beta^T x) \quad h(t; z_2) = h_0(t) \exp(\beta_1 z_2 + \beta^T x)$$

$z_1, z_2$ でのハザード比

$$\begin{aligned} \frac{h(t; z_1)}{h(t; z_2)} &= \frac{h_0(t) \exp(\beta_1 z_1 + \beta^T x)}{h_0(t) \exp(\beta_1 z_2 + \beta^T x)} = \frac{\exp(\beta_1 z_1 + \beta^T x)}{\exp(\beta_1 z_2 + \beta^T x)} \\ &= \exp(\beta_1 z_1 + \beta^T x - \beta_1 z_2 - \beta^T x) \\ &= \exp(\beta_1 (z_1 - z_2)) \end{aligned}$$

生存時間に関わらずハザード比は一定となる。すなわち、比例ハザードモデルで仮定されている比例性とは「異なる共変量を有する個体のハザードがすべての時間上で一定である」ことを意味する。

## 比例ハザードモデルとハザード比

治療間の違いのように、カテゴリカル変数を数値変数にしたものをダミー変数という。

ダミー変数として処理群1, 対照群0とおいた場合は

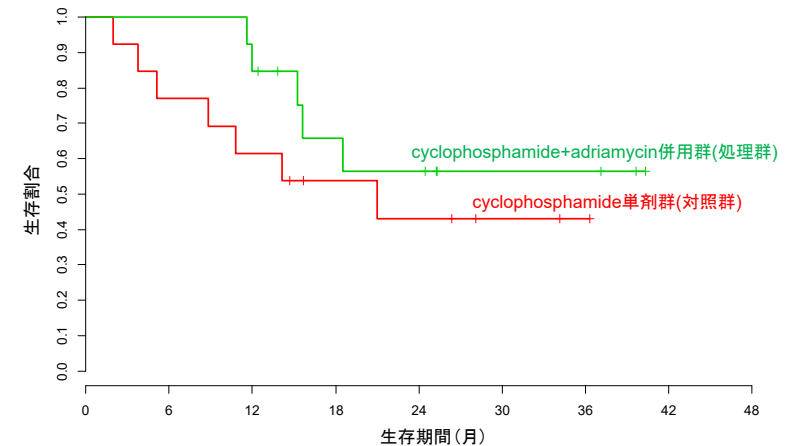
$$\begin{aligned} h(t; 0) &= h_0(t) \exp(\beta_1 \times 0 + \beta^T x) & h(t; 1) &= h_0(t) \exp(\beta_1 \times 1 + \beta^T x) \\ &= h_0(t) \exp(\beta^T x) & &= h_0(t) \exp(\beta_1 + \beta^T x) \end{aligned}$$

ハザード比

$$\begin{aligned} \frac{h(t; 1)}{h(t; 0)} &= \frac{h_0(t) \exp(\beta_1 + \beta^T x)}{h_0(t) \exp(\beta^T x)} = \frac{\exp(\beta_1 + \beta^T x)}{\exp(\beta^T x)} \\ &= \exp(\beta_1 + \beta^T x - \beta^T x) \\ &= \exp(\beta_1) \end{aligned}$$

すなわち、回帰係数の指数値は、2群でのハザード比を表す。これを用いることで、例えば、回帰係数に対する信頼区間に基づいて、ハザード比の信頼区間を計算できる。

例示: ECOG (Eastern Cooperative Oncology Group) が実施した卵巣癌に対するランダム化比較試験のデータ (Edmunson et al., 1979)



いま、処理群を1, 対照群を0としたときの比例ハザードモデルは

$$h(t; x) = h_0(t) \exp(-0.595 \cdot x_{\text{treat}})$$

なので、ハザード比HRは、次のように計算できる。

$$HR = \exp(-0.595) = 0.551$$

## 回帰係数に対する3種類の検定

比例ハザードモデルの回帰係数 $\beta$ の検定には、3種類が存在する。

**Wald検定, スコア検定, 尤度比検定**

いずれの検定においても、仮説は同じである。

帰無仮説 $H_0$  : 回帰係数 $\beta$ は0である。

対立仮説 $H_1$  : 回帰係数 $\beta$ は0でない。

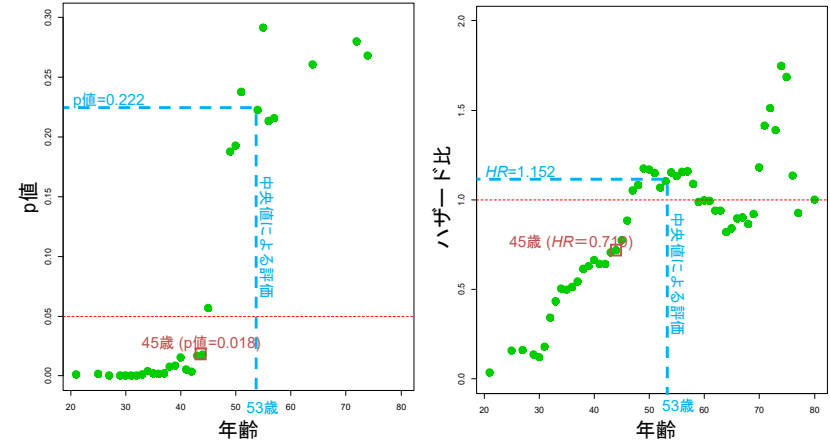
3種類の検定では、仮説に対するアプローチが違うが、どれを用いても誤りではない(因みに、治療をダミー変数で表した比例ハザードモデルの回帰係数のスコア検定のp値とログランク検定のp値は一致する)。

ECOG(Eastern Cooperative Oncology Group)が実施した卵巣癌に対するランダム化比較試験のデータ(Edmunson et al., 1979)

検定方法	p値
尤度比検定	0.305
Wald検定	0.310
スコア検定	0.303

## ちょっと脱線：共変量が量的尺度の場合の2値化の留意点

ドイツ乳がん研究グループによって実施されたランダム化比較試験のデータ(Schumacher et al., 1994)



変量の2値化において用いられることが多い中央値(53歳以上, 53歳未満)をカットオフ値にした場合, p値は0.222である(有意でなかった). 一方で, 45歳(45歳以上, 45歳未満)をカットオフ値にした場合, p値は0.018である(有意だった).

Schumacher, M., et al. : Randomized 2x2 trial evaluating hormonal treatment and the duration of chemotherapy in node-positive breast cancer patients. Journal of Clinical Oncology, 12, 2086-2093, 1994.

## ちょっと脱線：共変量が量的尺度の場合の2値化の留意点

**45歳(<45/≥45)の場合**

高齢群(45歳以上)のほうが低年齢群(45歳未満)よりも死亡リスクが**低い** (HR=0.719)

解釈が逆

**53歳(<53/≥53)の場合**

高齢群(53歳以上)のほうが低年齢群(53歳未満)よりも死亡リスクが**低い** (HR=1.152)

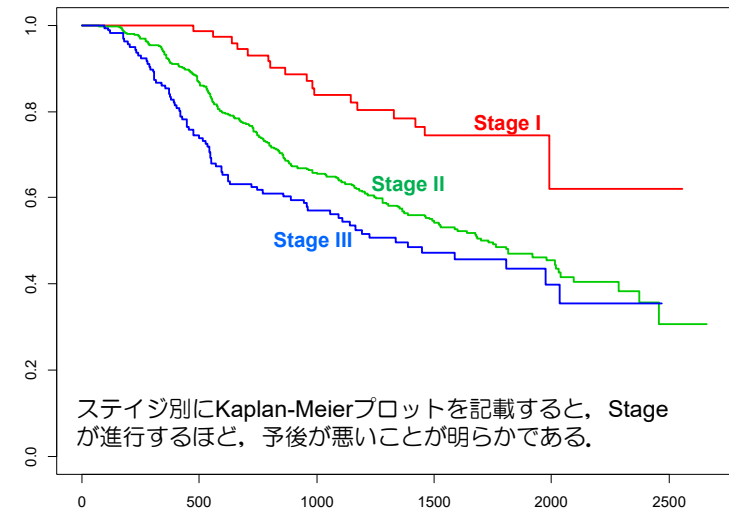
## カットオフ値の選定に関するRule of Thumb

- (1) 中央値を用いる
- (2) 臨床的にリーズナブルなカットオフ値を選定する

臨床的にリーズナブルとは、今回の事例の場合には、例えば疫学的調査などから得られた乳癌患者の平均年齢を用いたり、あるいは、閉経年齢を用いることを意味する。いずれにしても、カットオフ値選定におけるゴールド・スタンダードは存在しないため、試行的に決定する必要がある。

## 順序カテゴリカル変数の取り扱い

ドイツで実施された、ホルモン療法の有無による有効性・安全性を評価するための無作為化比較第III相試験の例



ステージ別にKaplan-Meierプロットを記載すると、Stageが進行するほど、予後が悪いことが明らかである。

## 順序カテゴリカル変数の取り扱い

### 取り扱い案1

$$\text{ダミー変数: } x_{\text{Stage}} = \begin{cases} 0, & \text{Stage I} \\ 1, & \text{Stage II} \\ 2, & \text{Stage III} \end{cases} \quad \text{比例ハザード・モデル: } h(t; x) = h_0(t) \exp(\beta \cdot x_{\text{Stage}})$$

この場合、Stage IとStage IIのハザード比、Stage IIとStage IIIのハザード比が同じであることを仮定するので、あまり用いられない。

### 取り扱い案2

$$\text{ダミー変数: } x_{\text{SII}} = \begin{cases} 1, & \text{Stage II} \\ 0, & \text{その他} \end{cases}, \quad x_{\text{SIII}} = \begin{cases} 1, & \text{Stage III} \\ 0, & \text{その他} \end{cases}$$

$$\text{比例ハザード・モデル: } h(t; x) = h_0(t) \exp(\beta_1 \cdot x_{\text{SII}} + \beta_2 \cdot x_{\text{SIII}})$$

	$x_{\text{SII}}$	$x_{\text{SIII}}$	
Stage I	0	0	← Reference
Stage II	1	0	← $\exp(\beta_1)$ はStage II/Stage IのHR
Stage III	0	1	← $\exp(\beta_2)$ はStage III/Stage IのHR

統計パッケージ(JMP等)では、順序カテゴリカルデータについては、こちら形式で取り扱っている。

## 順序カテゴリカル変数の取り扱い



いまの事例では、ダミー変数を2個使うので、回帰係数も2個になる。その場合には、Stageが予後に影響するかどうかを検定できないのでは？

$$h(t; x) = h_0(t) \exp(\beta_1 \cdot x_{\text{SII}} + \beta_2 \cdot x_{\text{SIII}})$$

Stage IIによる影響    Stage IIIによる影響

$\beta_1, \beta_2$ のいずれかに意味があれば、Stageの影響があるといえる

先ほどの3種類の検定は複数の回帰係数にまたがって行うことができる。つまり、Stageの場合を例にとると、仮説は次の通り：

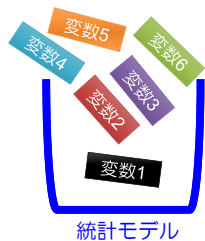
帰無仮説 $H_0$  :  $\beta_1 = \beta_2 = 0$ である(すべての回帰係数は0である)。  
対立仮説 $H_1$  :  $H_0$ ではない。



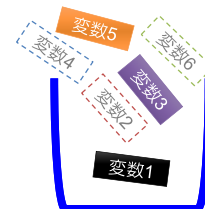
	ハザード比	p値
Stage II	2.39 (1.48-3.87)	<0.001
Stage III	3.17 (1.90-5.29)	<0.001

全体のWald検定の結果 : p値 <0.001

## 変数選択の動機



多変量解析では、たくさんの説明変数(共変量)を入れるほど情報がたくさんになるので、良い統計モデルになるということ？

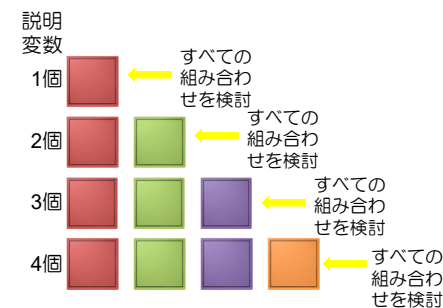


必ずしもそうではない。説明変数が多いほど不必要な変数は単なるノイズでしかないわけだし、また、**多重共線性**の問題などがある。そのため、必要なものだけで統計モデルをつくるのが推奨されることが多い。そのための方法が**変数選択**である。



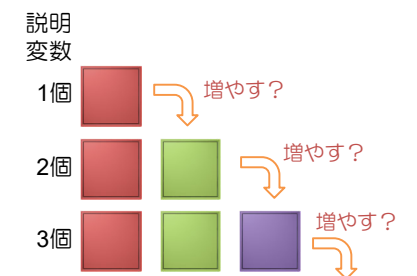
## 変数選択のアルゴリズムの種類

### 総当り法



説明変数が1個, 2個, ...のそれぞれのパターンを計算しその中から最適なものを選ぶ

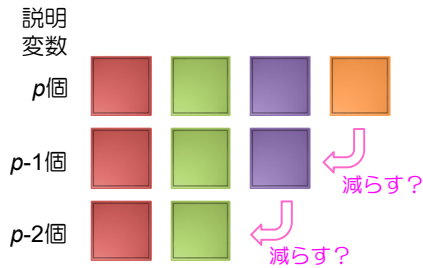
### 変数増加法 (前進ステップワイズ法)



説明変数が1個の場合からスタートして、変数を追加したほうが良ければ増やし、そうでなければ変数の追加をしない。

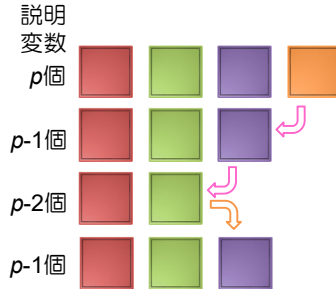
## 変数選択のアルゴリズムの種類

### 変数減少法 (後退ステップワイズ法)



全ての説明変数からスタートして、変数を減らしても影響がなければ減らし、そうでなければ変数の削除をしない。

### 変数増減法 (ステップワイズ法)



変数減少法からスタートするが、変数増減法では変数が削除された場合にその変数を考慮しなかったが、変数増減法では変数の削除と削除した変数の追加の両方を検討しながら各ステップを進める。

## 変数選択における経験則

1. 評価したい要因は変数選択に強制的に加える。
2. 変数増加法の落とし穴  
標本サイズが少ない場合には、本来必要な変数を追加する前にアルゴリズムが終了する「取りこぼし」が発生する
3. 多数の共変量がある場合の留意点  
多数の共変量がある場合には、事前スクリーニングとして単回帰分析などでスクリーニングを行うことが推奨される(ただし、p値を0.05にする必要は必ずしもない)。
4. 欠測が多い共変量は、欠測が多い理由(例えば、検査が煩雑でありあまり行われていないなど)があることが多い。また、欠測があった被験者が除外されるので、共変量自体を除外するほうがよい。
5. 可能であれば総当たり法を用いる(JMPでは不可能)

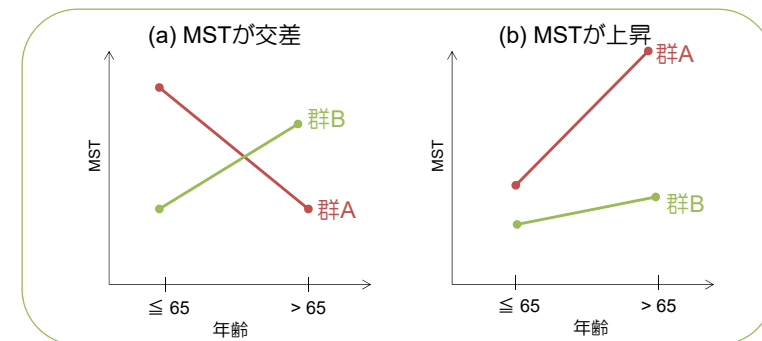
## 乳がんデータの解析例

	単変量			多変量		
	回帰係数	ハザード比	p値	回帰係数	ハザード比	p値
ホルモン療法の有無(有/無)	-0.364	0.125	0.004	-0.436	0.646	0.001
閉経の有無(無/有)	-0.063	0.939	0.596	0.169	1.184	0.397
グレード(II/I)	0.872	2.391	<0.001	0.520	1.683	0.040
(III/I)	1.154	3.170	<0.001	0.559	1.748	0.044
年齢(>53/≤53)	0.100	1.105	0.390	0.304	1.356	0.116
腫瘍径(>25/≤25)	0.325	1.384	0.005	0.109	1.116	0.364
転移個数(>3/≤3)	0.945	2.573	<0.001	0.914	2.494	<0.001
プロゲステロンレセプター(>32/≤32)	-0.764	0.466	<0.001	-0.673	0.510	<0.001
エストロゲンレセプター(>36/≥36)	-0.326	0.722	0.005	0.065	1.067	0.627

	変数選択: AIC			変数選択: BIC		
	回帰係数	ハザード比	p値	回帰係数	ハザード比	p値
ホルモン療法の有無(有/無)	-0.443	0.642	0.001	-0.413	0.662	0.001
閉経の有無(無/有)	—	—	—	—	—	—
グレード(II/I)	0.509	1.663	0.044	—	—	—
(III/I)	0.564	1.758	0.041	—	—	—
年齢(>53/≤53)	0.181	1.199	0.125	—	—	—
腫瘍径(>25/≤25)	—	—	—	—	—	—
転移個数(>3/≤3)	0.929	2.532	<0.001	0.951	2.589	<0.001
プロゲステロンレセプター(>32/≤32)	-0.641	0.527	<0.001	-0.735	0.480	<0.001
エストロゲンレセプター(>36/≥36)	—	—	—	—	—	—

## 交互作用

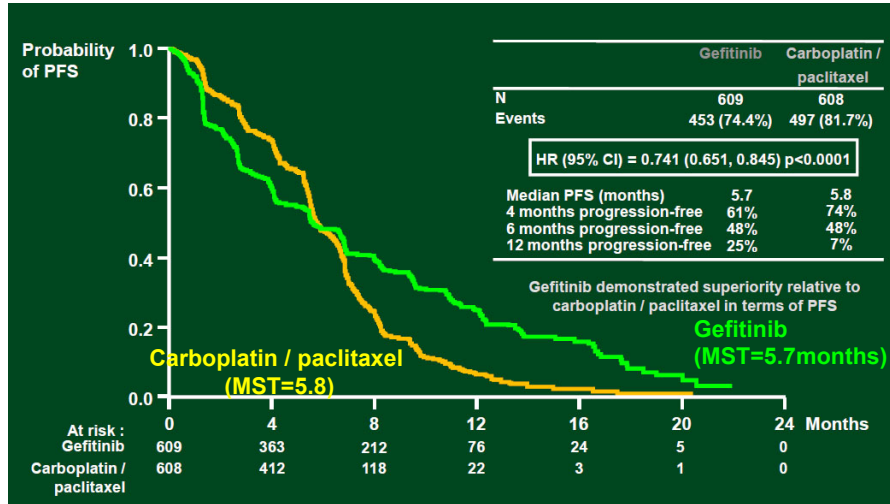
交互作用の例：下記は抗がん剤Aと抗がん剤Bでの中央生存時間(MST)をプロットしたものである。



何らかの予後因子によって、抗がん剤の効果は逆転したり、相乗効果によって、より効果が上昇するような現象を交互作用(interaction)という。

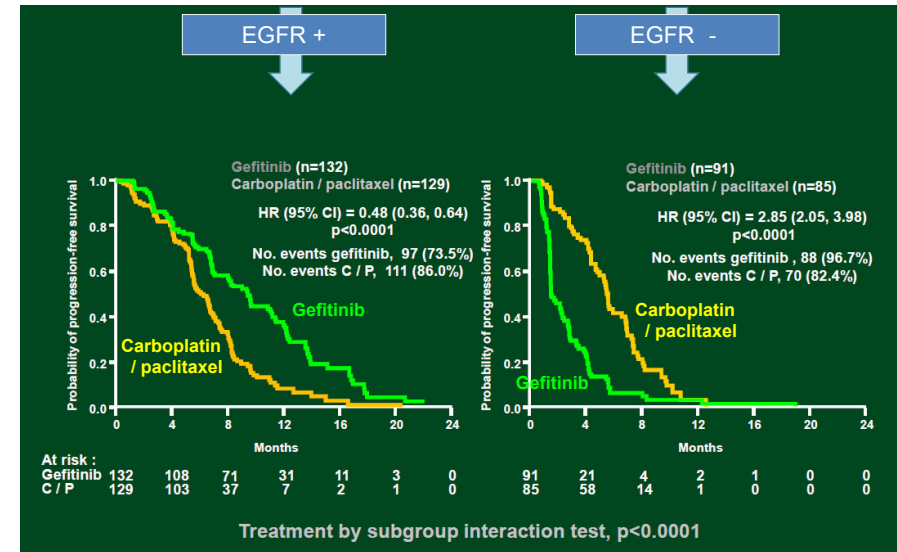
# 交互作用の例示：IPASS Study

## IPASS STUDY



Mok, T.S. et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma, N.Engl.J Med., 361(10), 947-57, 2009.

## EGFRで分けると...



Mok, T.S. et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma, N.Engl.J Med., 361(10), 947-57, 2009.

## 交絡と交互作用

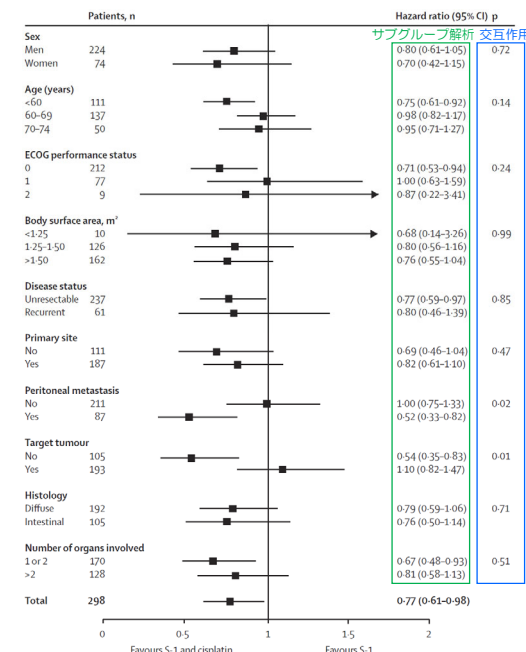
5. 第III相試験

交絡	交互作用
<ul style="list-style-type: none"> <li>重要な予後因子が群間で不均衡</li> <li>試験計画そのものが問題である (割付因子として組み込む必要性)</li> <li>交絡がないということは<b>試験計画の妥当性</b>がいえる。</li> </ul>	<ul style="list-style-type: none"> <li>薬剤の効果が、サブグループ間で異なる。</li> <li>薬剤の適応患者像に関する制限につながる。</li> <li>交互作用がないということは<b>治療の一般可能性</b>のあることがいえる。</li> </ul>

「浜田知久馬(1999). 学会・論文発表のための統計学, 新興交医書出版部」より引用・改編

- (1) ICH E5ガイドライン「臨床試験の統計的原則」では、多施設共同試験では、薬剤と施設の交互作用を検討することが要求される。これは、薬剤の一般可能性を証明するためである。
- (2) 交互作用は、回帰モデル(生存時間解析では、比例ハザードモデル)に治療(薬剤)のダミー変数と予後因子の積による項を追加し、その回帰係数の有意性により検出する。一方で、交互作用の構成については、Cox(1984)によって詳細に吟味されている。

Cox, D.R.(1984). Interaction, *International Statistical Review*, 52(1), 1-31.



## 実際の文献事例：SPIRITS trial

進行・再発胃癌患者に対する1st lineとしての、S-1単独療法とS-1+CDDP療法を比較した無作為化比較第3相試験

### サブセット解析

ハザード比の95%信頼区間が1.0をまたがない場合には、そのサブセットにおいて治療効果に違いがあることを示す。

### 交互作用の評価

有意である(0.05を下回る)場合には、当該要因によって、治療効果に違いが生じることを意味する。

Koizumi, W. et al.: S-1 plus cisplatin versus S-1 alone for first-line treatment of advanced gastric cancer (SPIRITS trial): a phase III trial, *Lancet Oncol.*, 9(3), 215-221, 2008.

## 予後因子(prognosis factor)と予測因子(predictive factor)

最近、バイオマーカーを伴う臨床試験も増えている。いま、共変量を

$x_{\text{treat}}$  : 新治療を1, 既存治療を0とするダミー変数  
 $x_{\text{marker}}$  : マーカー陽性を1, 陰性を0とするダミー変数

のように定義した場合の比例ハザードモデルは下式になる:

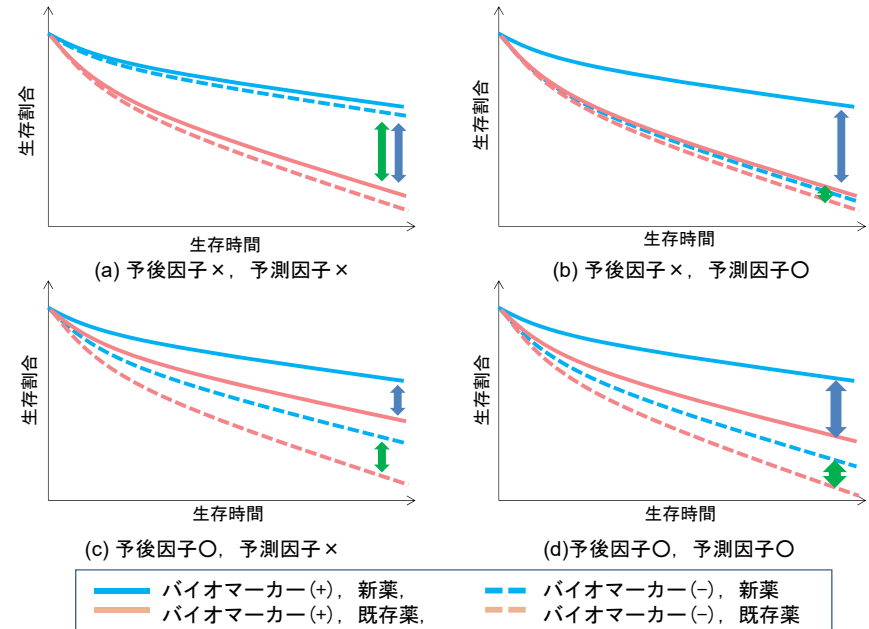
$$h(t; x) = h_0(t) \cdot \exp \{ \beta_1 \cdot x_{\text{treat}} + \beta_2 \cdot x_{\text{marker}} + \beta_3 \cdot x_{\text{treat}} \cdot x_{\text{marker}} \}$$

治療の違いによる影響を表す  
(治療効果).

バイオマーカーによる影響を表す  
(予後因子).

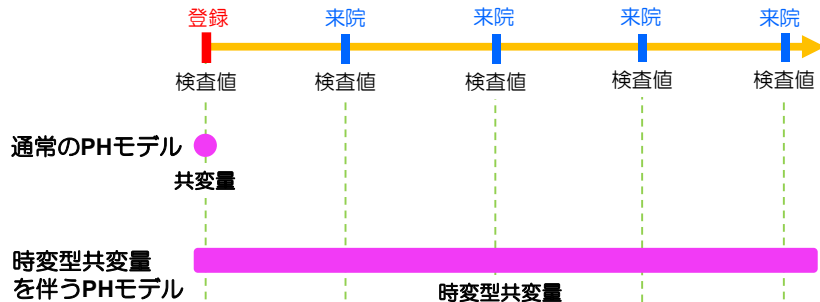
バイオマーカーによる治療効果の違いを表す  
(予測因子).

すなわち、予測因子とは、治療との交互作用で表される。



## 特殊な状況での比例ハザードモデル(1)

### ■ 時変型共変量を伴う比例ハザードモデル



- 臨床試験では、登録時だけでなく、来院時等においてもデータが取得される。
- 一方で、通常の比例ハザードモデルの共変量には登録時の検査データを用いるものの、来院時等でのデータを共変量(多変量解析)に用いない。
- 時変型共変量を伴う比例ハザードモデルとは時間経過とともに変化する共変量(時変型共変量)を取り扱う。
- ただし、時変型共変量は解釈が難しいため、あまり用いられない。

## 特殊な状況での比例ハザードモデル(2)

### ■ 層別比例ハザードモデル

#### 通常の比例ハザード・モデル

$$\text{ハザード} = \text{基線ハザード (BH)} \times \exp \left( \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \right)$$

時間*t*のみに依存

共変量: 治療法, 性別, 腫瘍径などのこと

係数1×共変量1 + 係数2×共変量2 + ...

共変量*x*のみに依存

例えば、国際共同試験などの場合には、国や民族差・人種差などが違うため、基線ハザード自体が異なる可能性がある

#### 層別比例ハザード・モデル

$$\text{ハザード} = \begin{matrix} \text{BH1} \\ \text{BH2} \\ \text{BH3} \end{matrix} \times \exp \left( \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \right)$$

時間*t*のみに依存

共変量: 治療法, 性別, 腫瘍径などのこと

係数1×共変量1 + 係数2×共変量2 + ...

共変量*x*のみに依存